DOCUMENT RESUME

ED 435 619 TM 030 216

AUTHOR Samejima, Fumiko
TITLE General Graded Response Model.
PUB DATE 1999-04-20
NOTE 39p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Montreal, Quebec, Canada, April 19-23, 1999).
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Equations (Mathematics); Estimation (Mathematics); *Mathematical Models; *Selection
IDENTIFIERS *Graded Response Model; Polytomous Items

ABSTRACT

This paper describes the graded response model. The graded response model represents a family of mathematical models that deal with ordered polytomous categories, such as: (1) letter grading; (2) an attitude survey with "strongly disagree, disagree, agree, and strongly agree" choices; (3) partial credit given in accord with an individual's degree of attainment; and (4) computerized cognitive diagnosis. Equations are given for the general model and for homogeneous and heterogeneous cases. Conditions are outlined for model selection, and estimation of the operating characteristics is described. (Contains 1 table, 10 figures, and 18 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made from the original document.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY
Fumiko Samejima
TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
✓ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to
improve reproduction quality.
• Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy.

ED 435 619

TM030216

GENERAL GRADED RESPONSE MODEL[1]

FUMIKO SAMEJIMA

UNIVERSITY OF TENNESSEE

The 1999 Annual NCME Meeting

April 20, 1999

Montreal, Canada

BEST COPY AVAILABLE

[1] Requests for reprints should be sent to Fumiko Samejima, Department of Psychology, 405 Austin Peay Bldg., University of Tennessee, Knoxville, Tennessee 37996-0900. E-mail: samejima@psych1.psych.utk.edu

Introduction

The *graded response model* represents a family of mathematical models which deal with *ordered polychotomous categories*, that include:

1. **letter grading, A, B, C, D and F, in students' performance evaluation;**
2. ***strongly disagree*, *disagree*, *agree* and *strongly agree* in an attitude survey;**
3. ***partial credit* given in accordance with an individual's degree of attainment toward a problem solution;**
4. ***computerized cognitive diagnosis*; etc.**

General Graded Response Model

• Framework

1. *Operating Characteristic* $P_{x_g}(\theta)$:

   Conditional probability, given θ , with which the individual with ability θ receives a score x_g , that is,

$$P_{x_g}(\theta) \equiv prob.[X_g = x_g \mid \theta] \ ,$$

θ : **latent trait (e.g., ability, attitude, etc.) that assumes any real number,**

g : **an item, the smallest unit of manifest entity for measuring** θ **,**

X_g : **a graded item response to item** g **, with** $x_g\ (= 0, 1, ..., m_g)$ **as its realization.**

2. *Response Pattern* V :

A sequence of X_g **for** $g = 1, 2, ..., n$ **, with its realization,** v **such that**

$$v \ = \ \{ \ x_1, \ x_2, \ ..., \ x_g, \ ..., \ x_n \ \}' \ .$$

3. *Local Independence* **(Lord & Novick, 1968):**

It is assumed that within any group of individuals all characterized by the same value of ability θ **the distributions of the item responses are all** *independent* **of each other, that leads to:**

$$P_v(\theta) \ \equiv \ prob.[V = v \mid \theta] \ = \ \prod_{x_g \in v} P_{x_g}(\theta) \ ,$$

$P_v(\theta)$: **conditional probability, given** θ **, for the response pattern** v **, also the** *likelihood function* $L(v \mid \theta)$ **for** $V = v$.

• General Model

1. *Processing Function* $M_{x_g}(\theta)$ **(Samejima, 1995, 1997):**

Joint conditional probability with which the individual completes the step x_g *successfully*, **under the conditions that:**

(a) the individual's ability level is θ **, and**

(b) the steps up to $(x_g - 1)$ **have already been completed successfully.**

Assume that $M_{x_g}(\theta)$ **is** *non-decreasing* **in** θ **, indicating that each item has some direct and positive significance to the ability measured.**

$$M_{x_g}(\theta) \begin{cases} = 1 & \text{for } x_g = 0 \\ = 0 & \text{for } x_g = m_g + 1 \ , \end{cases} \tag{1}$$

for all θ **, indicating:**

i. everyone can *at least* **obtain the item score** 0 **, and**

ii. *no one* **is able to obtain the item score** $(m_g + 1)$ **.**

2. *Fundamental Formula* **:**

$$P_{x_g}(\theta) = \prod_{s \le x_g} M_s(\theta) \, [1 - M_{(x_g+1)}(\theta)] \ . \tag{2}$$

(Samejima, 1972).

3. *Cumulative Operating Characteristic* $P^*_{x_g}(\theta)$ **(Samejima, 1995):**

Conditional probability with which the individual of ability θ **completes the cognitive process successfully** *up to* **the step** x_g **or** *further*. **Thus**

$$P^*_{x_g}(\theta) = \prod_{s \leq x_g} M_s(\theta) \quad . \tag{3}$$

From Eqs. (2) and (3):

$$P_{x_g}(\theta) = P^*_{x_g}(\theta) - P^*_{(x_g+1)}(\theta) \quad . \tag{4}$$

$P^*_{x_g}(\theta)$ **becomes the item characteristic curve (ICC),** $P_g(\theta)$ **, when** $m_g = 1$ **, i.e., in the general** *dichotomous* **response model.** $P^*_{x_g}(\theta)$ **is also** *non-decreasing* **in** θ **, and from Eqs. (1) and (3)**

$$P^*_{x_g}(\theta) \begin{cases} = 1 & \text{for } x_g = 0 \\ = 0 & \text{for } x_g = m_g + 1 \end{cases}$$

for all θ **.**

An *alternative* **interpretation of Eq. (3) is that, assuming that the factors affecting the individual's attitude toward** x_g **can be classified into two distinct tendencies:**

(a) being *tentatively* **attracted by** x_g **, and**

(b) its simultaneous or later *rejection* **(Samejima, 1972).**

Homogeneous and Heterogeneous Cases

• Homogeneous Case

1. **Family of models in which** $P^*_{x_g}(\theta)$ **'s for** $x_g = 1, 2, ..., m_g$ **are** *identical* **in shape.**

   They are positioned alongside the *abscissa* **in accordance with the item score** x_g **.**

   FIGURE 1: Example of a set of $P^*_{x_g}(\theta)$ **in the homogeneous case.**

2. **General Formula:**

$$P^*_{x_g}(\theta) = \int_{-\infty}^{a_g(\theta - b_{x_g})} \psi(t)\, dt \quad , \tag{5}$$

where

$$-\infty = b_0 < b_1 < b_2 < ..... < b_{m_g} < b_{m_g+1} = \infty \quad ,$$

$\psi(\bullet)$ **: some density function.**

a_g **: item** *discrimination* **parameter.**

b_g **: item score** *difficulty* **parameter.**

3. *Examples* of Specific Models

(a) Normal ogive model (Samejima, 1969, 1972):

$$P_{x_g}(\theta) = \frac{1}{[2\pi]^{1/2}} \int_{a_g(\theta - b_{x_g+1})}^{a_g(\theta - b_{x_g})} \exp\left[\frac{-t^2}{2}\right] dt \quad . \tag{6}$$

FIGURE 2: Typical operating characteristics, $P_{x_g}(\theta)$.

FIGURE 3: $M_{x_g}(\theta)$'s and $P^*_{x_g}(\theta)$'s in the normal ogive model.

(b) Logistic Model (Samejima, 1969, 1972):

$$P_{x_g}(\theta) = \frac{\exp\left[-Da_g(\theta - b_{x_g+1})\right] - \exp\left[-Da_g(\theta - b_{x_g})\right]}{[1 + \exp\left[-Da_g(\theta - b_{x_g})\right]][1 + \exp\left[-Da_g(\theta - b_{x_g+1})\right]]} \quad , \tag{7}$$

D : scaling factor, usually 1.7 .

Rasch model is a special case of Eq. (7) when $Da_g = 1$.

FIGURE 4: $M_{x_g}(\theta)$ and $P^*_{x_g}(\theta)$ in the logistic model.

NOTE: Birnbaum (1968) proposed the logistic model for dichotomous responses as a *substitute* for the normal ogive model. On the *graded* response level, however, these two models are substantially different in $M_{x_g}(\theta)$.

• Heterogeneous Case

1. All models in which *not all* $P^*_{x_g}(\theta)$'s are identical in shape.

2. *Examples* of Specific Models

(a) Extended Bock's *nominal* model:

Bock's nominal response model (Bock, 1972):

$$P_{h_g}(\theta) = \frac{\exp[\alpha_{h_g}\,\theta + \beta_{h_g}]}{\sum_{s\in H_g}\exp[\alpha_s\,\theta + \beta_s]} \quad . \tag{8}$$

h_g **: a nominal response to item** g **.**

H_g **: the set of all** h_g **'s .**

$\alpha_{h_g}(>0)$, β_{h_g} **: item response parameters.**

Samejima (1972) demonstrated that Bock's nominal response model can be considered as a *graded* response model in the heterogeneous case, if h_g **in Eq. (8) is replaced by** x_g **, and**

$$\alpha_0 \leq \alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_{m_g} \quad , \tag{9}$$

where a *strict* inequality should hold, at least, at one place.

Samejima did *not* pursue this further, for the reason that this model is based on the assumption that the *conditional ratio*, given θ **, of the probabilities of any two discrete**

responses to item g is *invariant* regardless of the set of alternatives selected from the answer space: the same assumption used in the *individual choice behavior* (Luce, 1959). This is *not* acceptable in graded response situations in general.

Later, however, Masters (1982) proposed his *partial credit* model and Muraki (1992) proposed his *generalized* partial credit model, both of which are special cases of Bock's nominal model that satisfy Eq. (9) with a strict inequality everywhere.

i. Masters' partial credit model:

$$\alpha_{x_g} = x_g + 1 \qquad \text{for} \quad x_g = 0, 1, ..., m_g \quad .$$

ii. Muraki's generalized partial credit model:

$$\alpha_{x_g} = (x_g + 1)\, a_g \qquad \text{for} \quad x_g = 0, 1. ..., m_g \quad .$$

(b) Logistic positive exponent family of models (LPEF) (Samejima, 1998b):

$$P^*_{x_g}(\theta) = [\Psi_g(\theta)]^{\xi_{x_g}} \quad , \tag{10}$$

where

$$\Psi_g(\theta) = \frac{1}{1 + \exp\,[-D\,\alpha_g(\theta - \beta_g)]} \quad . \tag{11}$$

$\xi_{x_g}(>0)$: **acceleration parameter.**

FIGURE 5: Examples of $P^*_{x_g}(\theta)$ **in LPEF.**

(c) Acceleration model (Samejima, 1995, 1997):

$$M_{x_g}(\theta) = [\Psi_{x_g}(\theta)]^{\xi_{x_g}} \quad . \tag{12}$$

ξ_{x_g} (>0) : *step* **acceleration parameter.**

A family of models in which $\Psi_{x_g}(\theta)$ **in Eq. (12) is specified by a** *strictly increasing*, **five times differentiable function of** θ **with zero and unity as its two asymptotes.**

E.G., Problem solving that requires a number of subprocesses before attaining the solution.

Graded item scores, or partial credits, 1 **through** m_g **, are assigned for the successful completion of separate observable steps.**

E.G.,

$$\Psi_{x_g}(\theta) = \frac{1}{1+\exp\,[-D\,\alpha_{x_g}(\theta-\beta_{x_g})]} \quad . \tag{13}$$

α_{x_g} (>0) : *step* **discrimination parameter.**

β_{x_g} : *step* **location parameter.**

Expanded LPEF for *cognitive diagnosis*, **etc.**

Model Selection

- **Desirable Features**

1. **The *principle* and the set of assumptions behind the model should agree with the *psychological reality* in question.**

   FIGURES 6 and 7: Similarities of the operating characteristics provided by two or more different models. (Masters' partial credit model vs. acceleration model.)

   Figures 6 and 7 exemplify the fact that *curve fitting* alone is *not* a sufficient model validation.

2. **Satisfaction of the *unique maximum condition* (Samejima, 1969, 1972).**

   This assures that the *likelihood function* of *any* response pattern consisting of such responses has a *unique* local or terminal maximum.

   FIGURE 8: Illustrative examples of likelihood functions that has a unique modal point and multi-modal points, respectively.

Basic Function:

$$A_{x_g}(\theta) \equiv \frac{\partial}{\partial\theta} \log P_{x_g}(\theta) = \sum_{s \le x_g} \frac{\partial}{\partial\theta} \log M_s(\theta) + \frac{\partial}{\partial\theta} \log [1 - M_{(x_g+1)}(\theta)] \ . \tag{14}$$

(a) $A_{x_g}(\theta)$ **is strictly** *decreasing* **in** θ **, and**

(b) its upper and lower *asymptotes* **are nonnegative and non-positive, respectively.**

FIGURE 9: Examples of a set of basic functions in the extended Bock model (Masters' model).

Item Response Information Function:

$$I_{x_g}(\theta) \equiv -\frac{\partial^2}{\partial\theta^2} \log P_{x_g}(\theta) \ . \tag{15}$$

Alternatively, the unique maximum condition is satisfied if $I_{x_g}(\theta)$ **is** *positive* **for all** θ **.**

3. The model should provide the *ordered modal points* **of the operating characteristics in accordance with the item scores.**

A *sufficient*, **though not necessary, condition:**

$$A_{(x_g-1)}(\theta) < A_{x_g}(\theta) \qquad for \quad x_g = 1, 2, ..., m_g \ ,$$

for all θ **.**

4. *Additivity* **of the operating characteristics (Samejima, 1995, 1997).**

The operating characteristics still belong to the *same* mathematical model under:

(a) *finer* recategorizations, and

(b) *combinings* of two or more categories together.

FIGURE 10: Example of additivity (acceleration model).

Graded item scores, or partial credits, are more or less *incidental.*

E.G. 1. Letter grades, A, B, C, D, and F, are combined to *pass-fail* grades.

E.G. 2. With the advancement of computer technologies, more abundant information can be obtained from the individual's performance in *computerized experiments*, as we proceed in research, and thus finer recategorizations of the whole cognitive process become possible.

The criterion (a) **leads to:**

5. *Generalizability of the model to a continuous model.*

TABLE 1: Evaluation of various models in terms of the above criteria.

Estimation of the Operating Characteristics

- **Parametric Estimation**

  1. **Multilog (Thissen, 1991):**

     Direct expansion of Bock and Atkin's (1981) EM solution of the *marginal likelihood equations* for dichotomous responses for Samejima's logistic model and Bock's nominal response model and its extentions to graded response models.

  2. **Parscale (Muraki and Bock, 1993):**

     Essentially the same EM algorithm for the above models.

- **Nonparametric Estimation and Parameterization**

  When the model has *more than two* item response parameters, as is the case with the LPEF, the acceleration model, etc., it is recommendable to use a *nonparametric* method of estimating the operating characteristics, such as Levine's (1984) and Samejima's (1998a), and then *parameterize* the outcomes, using a very general semiparametric method, such as Ramsay and Wang's (1993).

This will ameliorate the problem of *indeterminancy* of the estimated item response parameters.

References

[1] Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord and M. R. Novick, *Statistical theories of mental test scores* (Chapters 17-20). Reading, MA: Addison Wesley.

[2] Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, *37*, 29-51.

[3] Bock, R. D. & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, *46*, 443-459.

[4] Levine, M. (1984). *An introduction to multilinear formula scoring theory* (Office of Naval Research Report, 84-4). Champaign, IL: Model-Based Measurement Laboratory, Education Building, University of Illinois.

[5] Lord, F. M. (1952). A theory of mental test scores. *Psychometric Monograph*, *No. 7*.

[6] Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading. MA: Addison Wesley.

[7] Luce, R. D. (1959). *Individual choice behavior*. New York, NY: Wiley.

[8] Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149-174.

[9] Muraki, E. (1992). A generalized partial credit model: application of an EM

algorithm. *Applied Psychological Measurement*, *16*, 159-176.

[10] Muraki, E. & Bock, R. D. (1993). *Parscale*. Chicago, IL: Scientific Software.

[11] Ramsay, J. O. & Wang, X. (1993). *Hybrid IRT models*. Paper presented at the meeting of the Psychometric Society, Berkeley, CA.

[12] Samejima, F. (1969). Estimation of ability using a response pattern of graded scores. *Psychometrika Monograph, No. 17.*

[13] Samejima, F. (1972). A general model for free-response data. *Psychometrika Monograph, No. 18.*

[14] Samejima, F. (1995). Acceleration model in the heterogeneous case of the general graded response model. *Psychometrika*, *60*, 549-572.

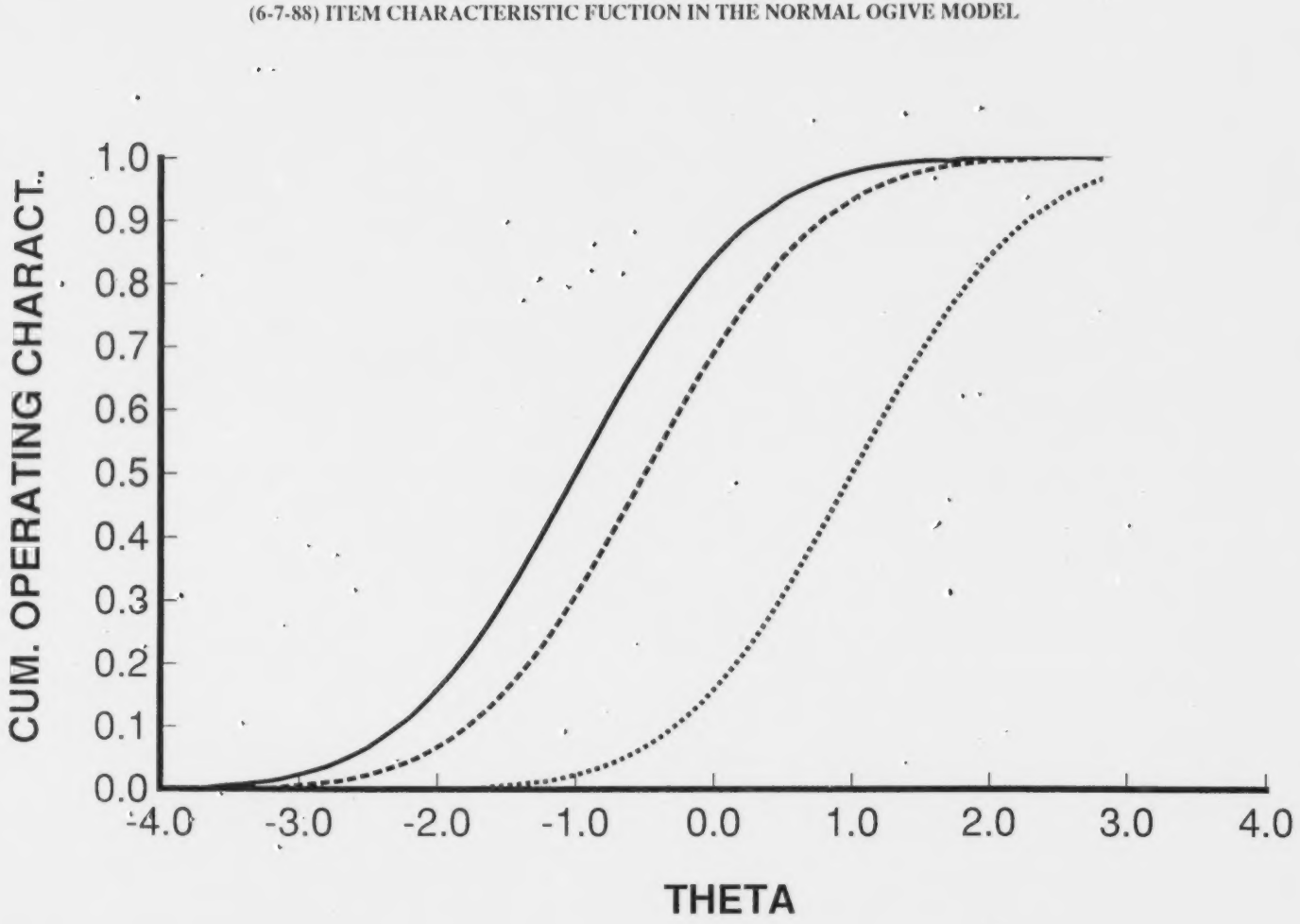
[15] Samejima, F. (1997). Graded response model. In W. J. van der Linden and R. K. Hambleton (Eds.), *Handbook of modern item response theory* (Chapter 5). New York, NY: Springer

[16] Samejima, F. (1998a). Efficient nonparametric approaches for estimating the operating characteristics of discrete item responses. *Psychometrika*, *63*, 111-130.

[17] Samejima, F. (1998b). The logistic positive exponent family of models: Virtue of asymmetric item characteristic curves. *Psychometrika*, (accepted).

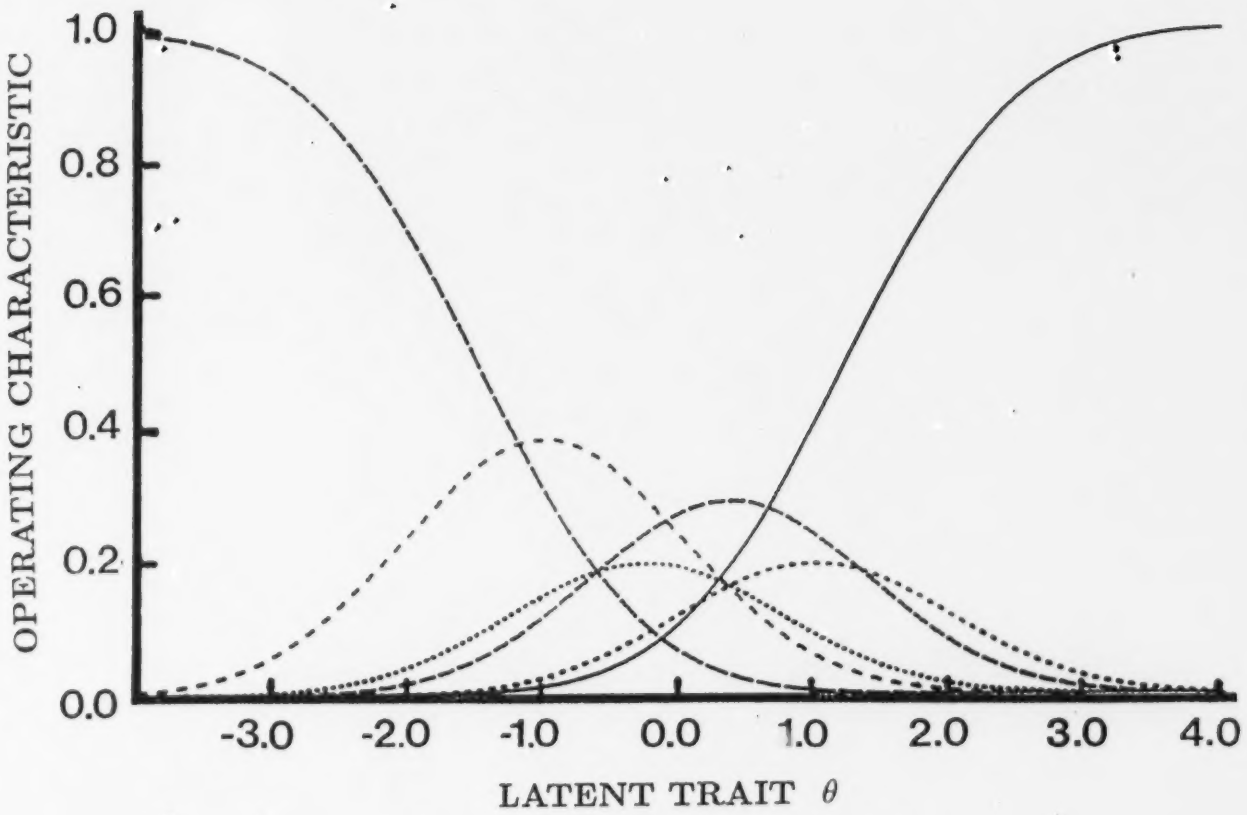
[18] Thissen D. (1991). *Multilog user's guide - version 6*. Chicago, IL: Scientific Software.

NCME99: Figure 1



0.561 0.50 1.50 6.50 8.00

NCME99: Figure 2

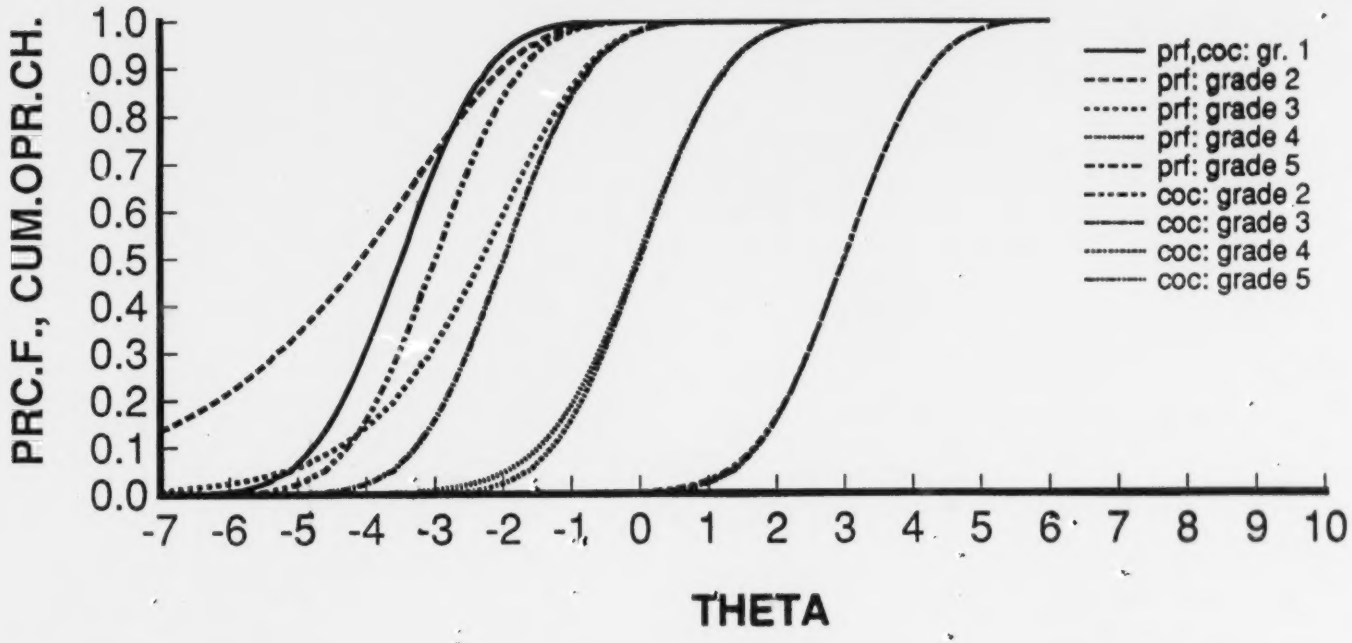


Normal Ogive Model; $a_g = 1.0$, $b_{x_g} = -1.50, -0.50, 0.00, 0.75, 1.25$

(Taken from ONR/RR-79-4: page 18)

SD~AERA95: 9504.F, PROCESSING FUNCTIONS, CUM.OP.CHARACTERISTI CS, SCORES 1-5, NML.OGV. MODEL, HOMOGENEOUS CASE, AG=1.0, BXG=-3.5,-3.0,-2.0,0 .0,3.0; 9504SFGR.RST2, 9504SFGR.RST3, 02/18/95

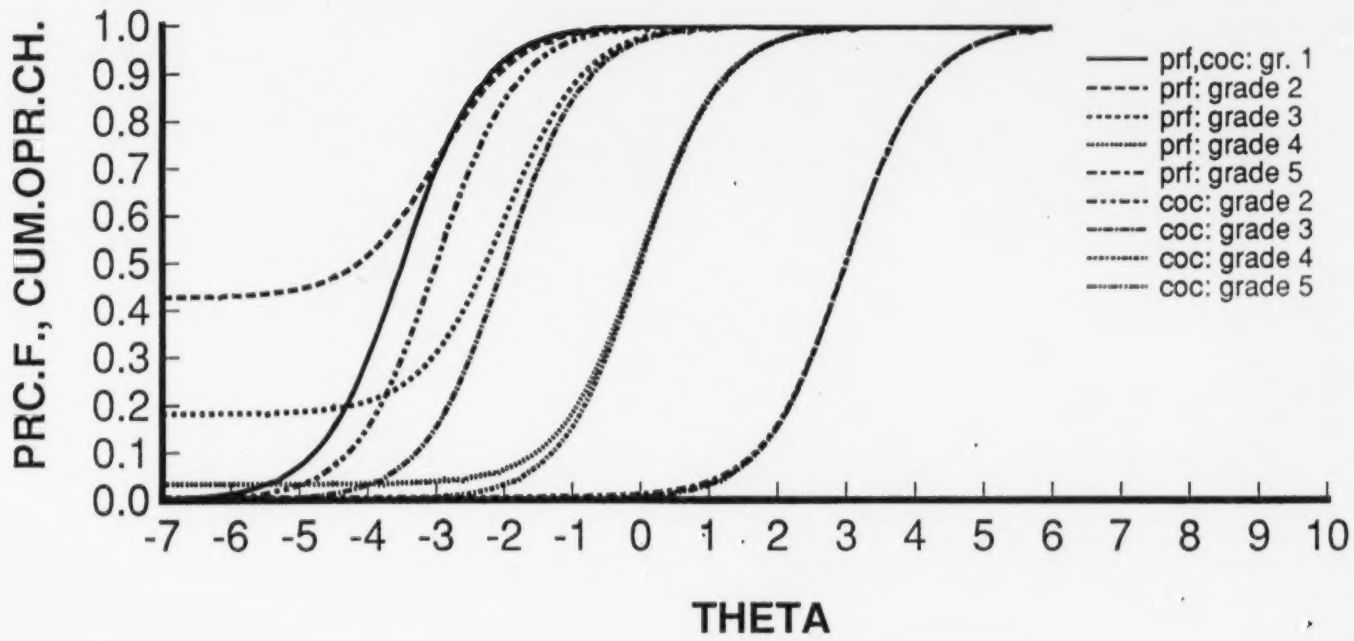
Normal Ogive Model



0.400 0.60 1.20 6.80 8.00
HOMOGNML.DAT, INHOMOG, plotted by F. Samejima

SD~AERA95: 9502.F, PROCESSING FUNCTIONS, CUM.OP.CHARACTERISTI CS, SCORES 1-5, LOGISTIC MODEL, HOMOGENEOUS CASE, AG=1.0, BXG=-3.5,-3.0,-2.0,0 .0,3.0; 9502SFGR.RST2, 9502SFGR.RST3, 02/16/95

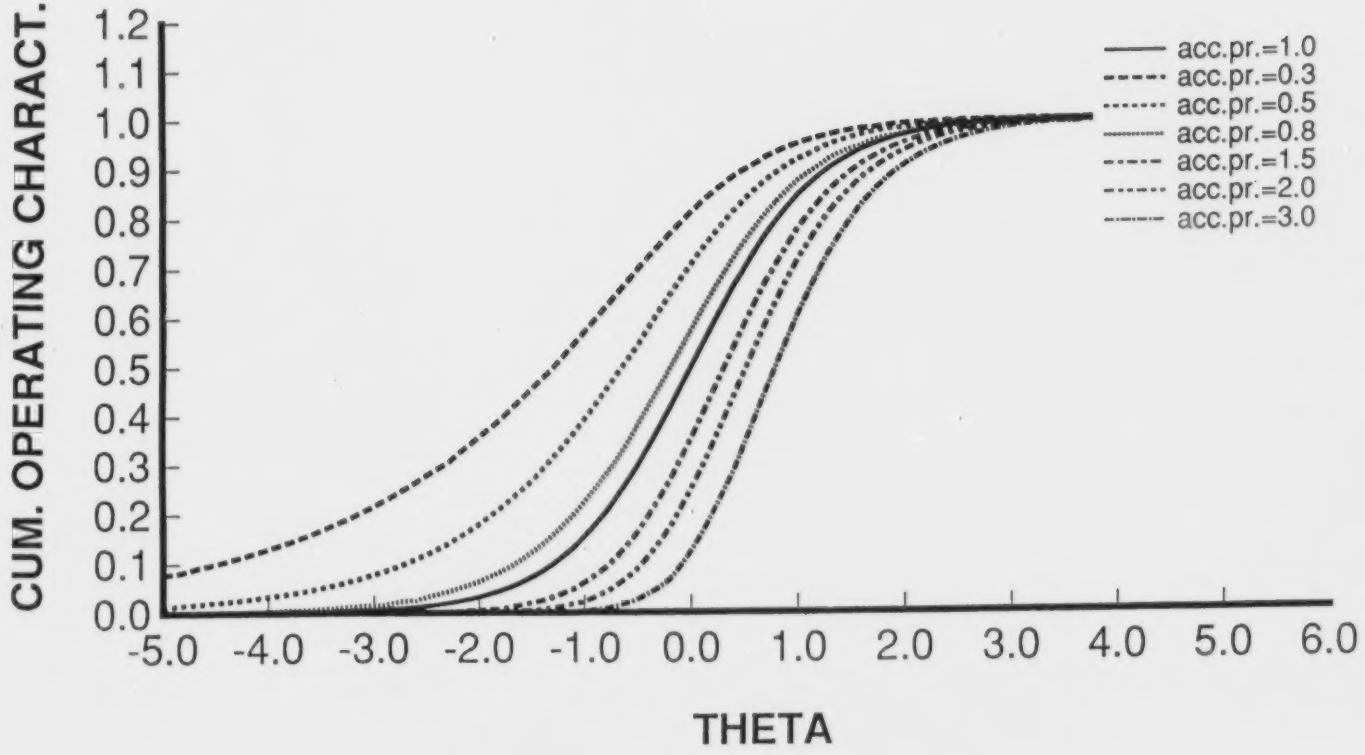
Logistic Model



0.400 0.60 1.20 6.80 8.00

HOMOGLOG.DAT, INHOMOG, plotted by F. Samejima

NCME99: Figure 5

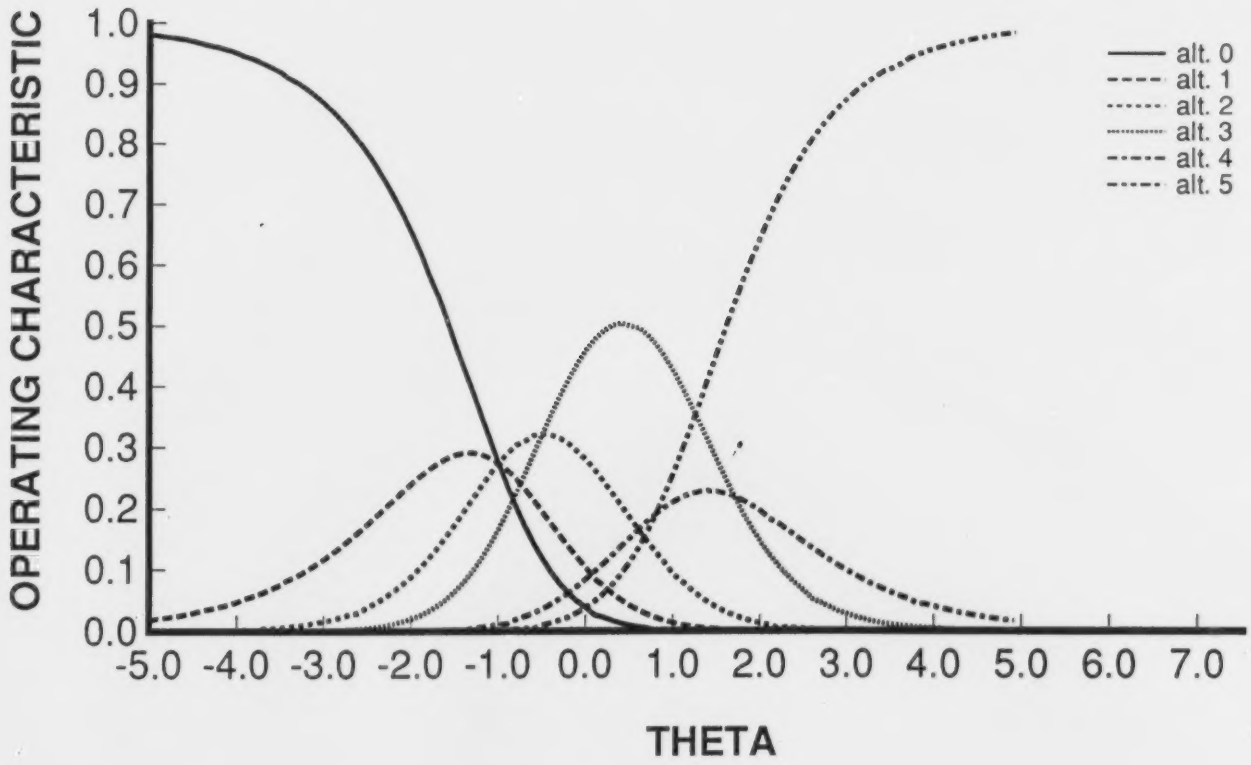


0.500 1.00 1.50 6.00 8.00

9601RSC1.DAT, IN9601IC, plotted by M. Foster

NCME99: Figure 6

SUN:SD~DBFS93:9303.F , BOCK.OP.CHAR., ITEM=YN08, GRADES 0-5, ALPHA=1,2,3,4,5,6, BETA=1,2,3,3.5,1.8,1 ,
9303BSYN4.RST2, 11/24/93



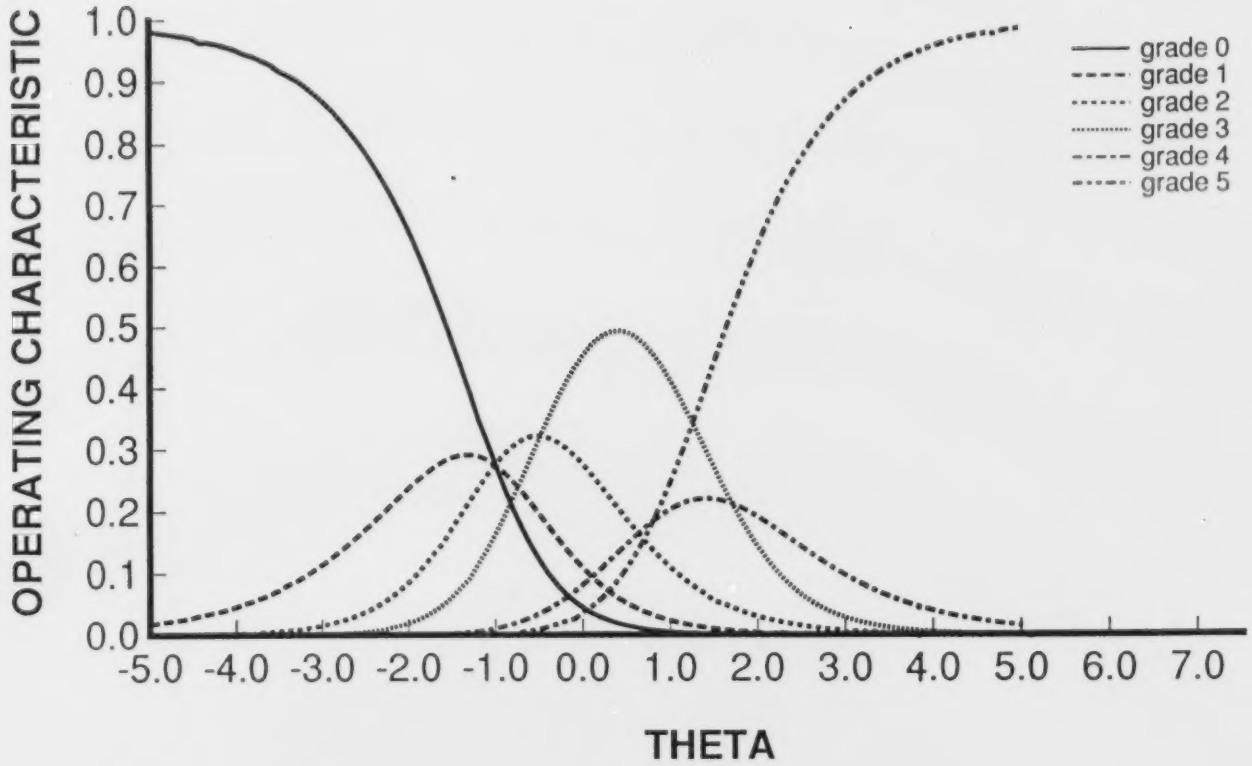
0.550 0.80 1.50 6.80 7.50

BSYN08B.DAT, INBFS93B, plotted by F. Samejima

23

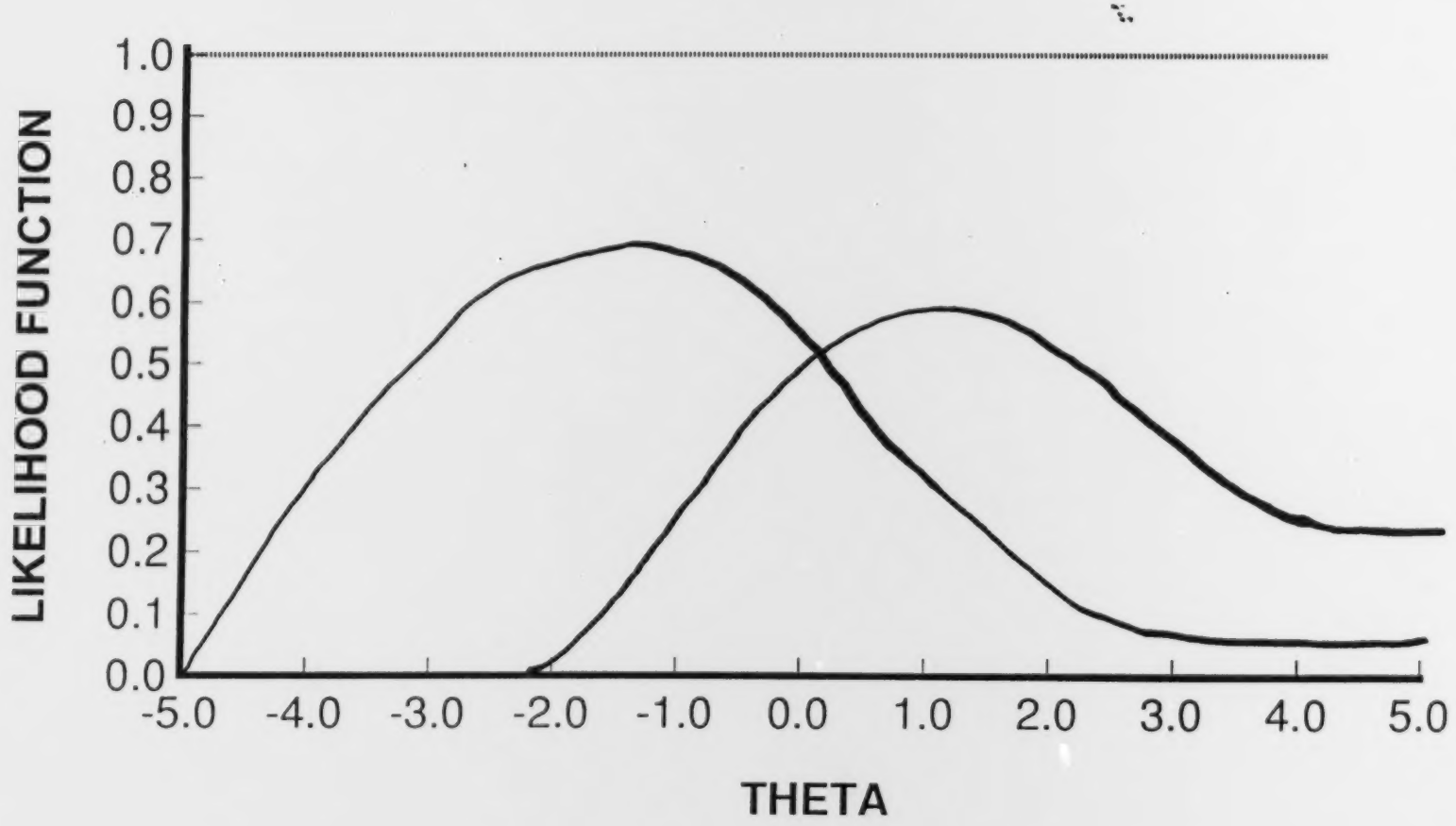
NCME99: Figure 7

HTRKS93: 9304.F, OPER.CHAR., ITEM=YN08, ALPHA=1,2,3,4,5,6, BETA=1,2,3,3.5,1.8, 1, PCR=0.3,0.5,0.7, 9304HTRKS3.RST1,
SIM:BOCK BY ACCEL. 11/30/93



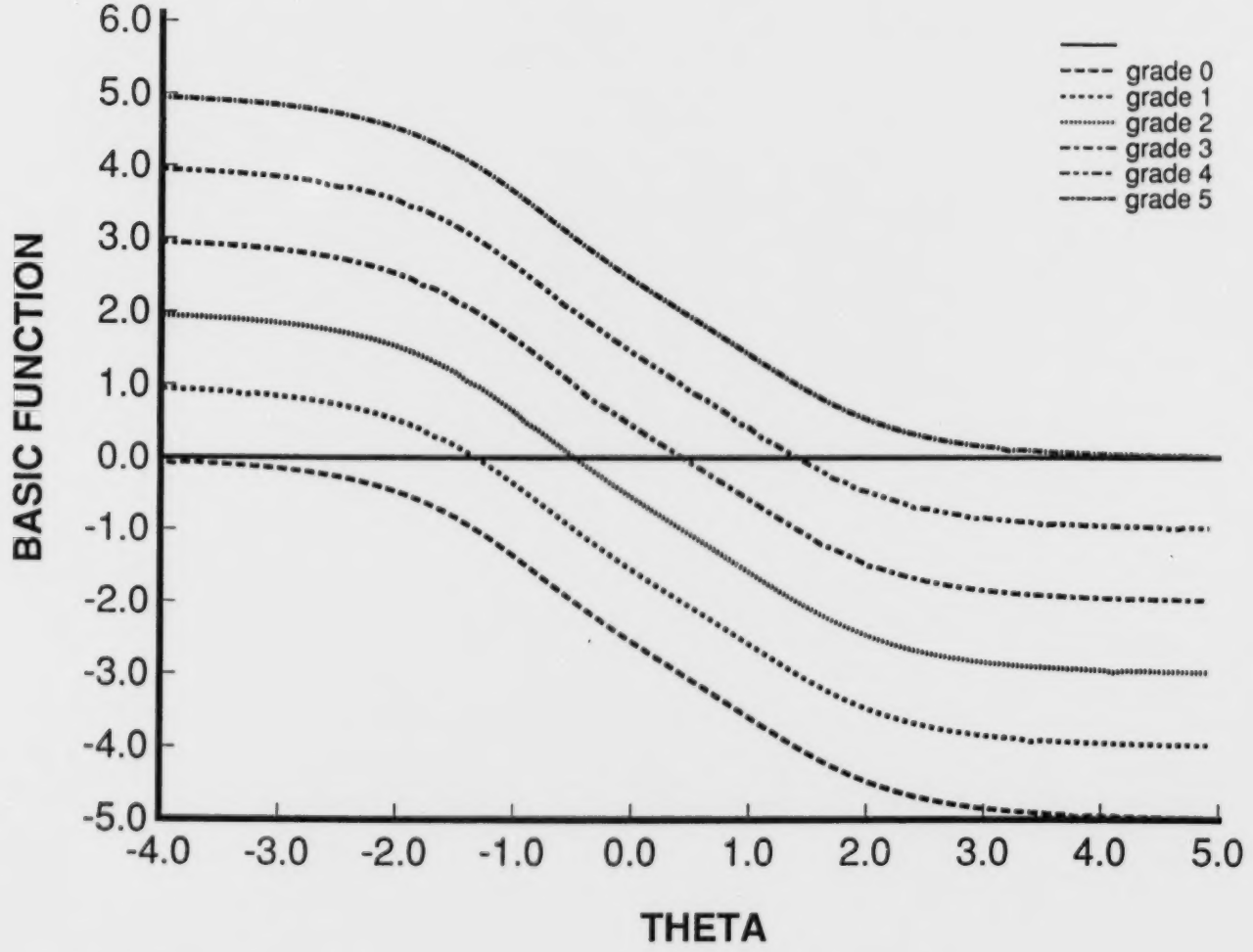
0.550 0.80 1.50 6.80 7.50
HT4YN83.DAT, INHTR93, plotted by F. Samejima

NCME99: Figure 8



NCME99: Figure 9

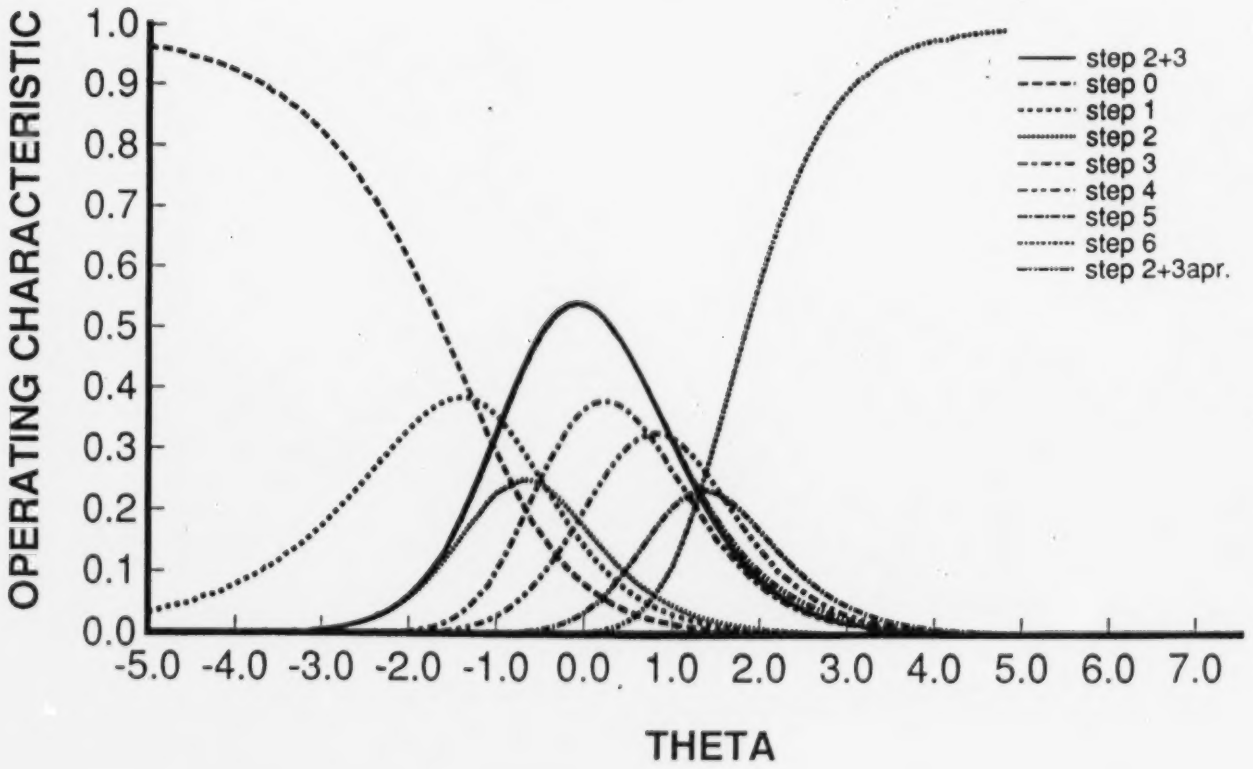
DBFS93: 9402.F, BASIC FUNCTION, ITEM=YN08, SCORES 0-5, ALPHA=1,2,3,4,5,6, BETA=1,2,3,3.5,1.8,1 , SUN:9402YN08.RST4, 02/10/94



0.750 0.80 1.50 6.80 7.50

DBFSYN8A.DAT, INHTR93A, plotted by F. Samejima

9302.F,SM9301.F; OPERATING CHARACTER.; ITEM=KS10,STEPS 0-6, STP.1,2,3:ALP.=1, BET.=-1,GAM.=.5,1,1. 5; STP.4,5,6:ALP.=1,BET .=1,GAM.=.5,1,1.5; STP.2+3, 2+3 APPR.; 9302ITMKS10MD.RST21, 9302HTRKS4.RST1, SM9301KS10.RST, 3/06/95



0.550 0.80 1.50 6.80 7.50

AERA95F7.DAT, INHTR93, plotted by F. Samejima

NCME99: Table 1

	Normal Ogive & Logistic Models	Acceleration Model	LPEF	Extended Bock Models
Additivity #1	Yes	Yes	Yes	No
#2	Yes	Robust	Yes	No
Generalizability to a continuous response model	Yes	Yes	Yes	No
Satisfaction of the unique maximum condition	Yes	Yes	Yes	Yes
Orderliness of modal points	Yes	Robust	Yes	Yes